

RESEARCH ARTICLE

In-depth proteomic characterization of *Schistosoma haematobium*: Towards the development of new tools for elimination

Javier Sotillo^{1,2*}, Mark S. Pearson¹, Luke Becker¹, Gebeyaw G. Mekonnen¹, Abena S. Amoah³, Govert van Dam³, Paul L. A. M. Corstjens⁴, Janice Murray⁵, Takafira Mduluzi^{6,7}, Francisca Mutapi^{5,7}, Alex Loukas^{1*}

1 Centre for Molecular Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, Queensland, Australia, **2** Laboratorio de Referencia en Parasitología, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain, **3** Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands, **4** Department of Molecular Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands, **5** Institute of Immunology & Infection Research, School of Biological Sciences, University of Edinburgh, Ashworth Laboratories, King's Buildings, Edinburgh, United Kingdom, **6** Biochemistry Department, University of Zimbabwe, Mount Pleasant, Harare, Zimbabwe, **7** TIBA Partnership, NIHR Global Health Research Unit Tackling Infections to Benefit Africa (TIBA), University of Zimbabwe

* javier.sotillo@isciii.es (JS); alex.loukas@jcu.edu.au (AL)



OPEN ACCESS

Citation: Sotillo J, Pearson MS, Becker L, Mekonnen GG, Amoah AS, van Dam G, et al. (2019) In-depth proteomic characterization of *Schistosoma haematobium*: Towards the development of new tools for elimination. PLoS Negl Trop Dis 13(5): e0007362. <https://doi.org/10.1371/journal.pntd.0007362>

Editor: Alessandra Morassutti, PUCRS, BRAZIL

Received: December 20, 2018

Accepted: April 5, 2019

Published: May 15, 2019

Copyright: © 2019 Sotillo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Mass spectrometry data along with the identification results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD011137 and 10.6019/PXD011137.

Funding: This work was supported by a program grant from the National Health and Medical Research Council (NHMRC - <https://nhmrc.gov.au/>) [program grant number 1037304] and a Senior Principal Research fellowship from NHMRC to AL (1117504). The funders had no role in the study

Abstract

Background

Schistosomiasis is a neglected disease affecting hundreds of millions worldwide. Of the three main species affecting humans, *Schistosoma haematobium* is the most common, and is the leading cause of urogenital schistosomiasis. *S. haematobium* infection can cause different urogenital clinical complications, particularly in the bladder, and furthermore, this parasite has been strongly linked with squamous cell carcinoma. A comprehensive analysis of the molecular composition of its different proteomes will contribute to developing new tools against this devastating disease.

Methods and findings

By combining a comprehensive protein fractionation approach consisting of OFFGEL electrophoresis with high-throughput mass spectrometry, we have performed the first in-depth characterisation of the different discrete proteomes of *S. haematobium* that are predicted to interact with human host tissues, including the secreted and tegumental proteomes of adult flukes and secreted and soluble egg proteomes. A total of 662, 239, 210 and 138 proteins were found in the adult tegument, adult secreted, soluble egg and secreted egg proteomes, respectively. In addition, we probed these distinct proteomes with urine to assess urinary antibody responses from naturally infected human subjects with different infection intensities, and identified adult fluke secreted and tegument extracts as being the best predictors of infection.

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

We provide a comprehensive dataset of proteins from the adult and egg stages of *S. haematobium* and highlight their utility as diagnostic markers of infection intensity. Protein composition was markedly different between the different extracts, highlighting the distinct subsets of proteins that different development stages present in their different niches. Furthermore, we have identified adult fluke ES and tegument extracts as best predictors of infection using urine antibodies of naturally infected people. This study provides the first steps towards the development of novel tools to control this important neglected tropical disease.

Author summary

Schistosomiasis is a neglected tropical disease affecting millions of people worldwide. Of the main three species affecting humans, *Schistosoma haematobium* is the most common, and is the leading cause of urogenital schistosomiasis. This parasite can cause a range of clinical complications associated with bladder pathogenesis, including squamous cell carcinoma as well as genital malignancy in women. Herein, we have performed the first comprehensive characterisation of the proteins implicated in host-parasite interactions (secreted and surface proteins from the adult flukes and secreted and soluble egg proteins) in order to advance our understanding of the parasite's biology. Furthermore, we have characterised the different antibody responses in urine from infected human subjects from an endemic area presenting different infection intensities. The data obtained in this study can be used as a first step towards the development of novel tools for the control of urogenital schistosomiasis.

Introduction

Schistosomiasis is a neglected tropical and debilitating disease caused by different trematodes from the genus *Schistosoma* [1]. It affects over 250 million people worldwide, particularly in developing and tropical regions [2–4]. Despite widespread use of the anthelmintic praziquantel in mass drug administration programs over the last 30 years [5], this parasitic infection still causes a loss of 1.9 million disability-adjusted life years (DALYs) [6], and this number could be greater if morbidity associated with asymptomatic infections was included in the calculations [7]. Of the 6 species affecting humans, *S. haematobium* is the most common, causing urogenital schistosomiasis in over 100 million people [1], although it is considered the neglected schistosome since the amount of “omics” information is scarce compared to other schistosomes and the difficulty in maintaining the parasite in an animal model [8, 9]. *S. haematobium* infection has been reported in 54 countries [10], particularly in sub-Saharan Africa and the Middle East [2, 3]. Furthermore, an outbreak of urogenital schistosomiasis was observed in Corsica (France) [11], although this parasite has recently been shown to be a hybrid between *S. haematobium* and *Schistosoma bovis*. [12].

The clinical complications associated with urogenital schistosomiasis are linked among others, to bladder pathology [13]. The association between squamous cell carcinoma and *S. haematobium* infection is undisputed [14]; indeed, this blood fluke has been classified as a Class I carcinogen by the International Agency for Research on Cancer (IARC) [15]. Once established in the mesenteric veins surrounding the bladder, *S. haematobium* adult females start laying

eggs that pass through the bladder epithelium to be secreted in the urine. However, some eggs get trapped in the bladder wall causing a chronic local inflammation that will develop into a granuloma accompanied by relentless cell proliferation and ultimately in some patients, bladder cancer [16] as well as genital malignancy in women [17, 18]. The initial inflammatory response is thought to be a reaction to mechanical damage caused by passing of eggs through the urothelium (a multilayered epithelium that lines most of the urogenital tract) [13]; however, proteins secreted by parasite eggs have also been shown to increase cell proliferation and angiogenesis [13]. One of the most abundant proteins secreted by *S. haematobium* eggs is the IPSE, or interleukin-4 inducing principle from *Schistosoma mansoni* eggs. IPSE induces cell proliferation and angiogenesis [13, 19], and an IPSE homologue from *S. haematobium* has been shown to attenuate inflammation by stimulating release of IL-4 [19, 20].

Despite the clear role of soluble egg proteins in the development of granulomas and a carcinogenic environment, the protein composition of this or other *S. haematobium* tissues/organs has not been fully characterised, although a few studies have identified several immunogenic proteins [21]. In contrast, different proteomes from the related species *S. mansoni*, *S. japonicum* and even from the cattle-infecting *S. bovis* have been well characterised [22–26], which has allowed for the identification of novel vaccine and diagnostic antigen candidates. Indeed, different tegumental and secreted proteins are potential vaccine candidates and some have entered or completed Phase I and II clinical trials [27]. The most progressed vaccine candidate for *S. haematobium* infection is glutathione S-transferase (Sh28GST), which completed Phase III clinical testing in 2012 [28] but the results have yet to be published [29].

Characterising the molecular interface of host-parasite interactions in *S. haematobium* infection is crucial for (i) a better understanding of the parasite's biology and, (ii) for the development of new tools for control and diagnosis. In the present study, we provide the first in-depth identification of secreted and surface proteomes of *S. haematobium*. We also characterise the antibody responses in urine from infected human subjects from an endemic area presenting different infection intensities as a first step towards the development of novel diagnostic tools against this devastating disease.

Results

***Schistosoma haematobium* tissue proteomes involved at the host-parasite interface**

The proteomes from different parasite extracts involved in host-parasite interactions, such as the adult ES and tegument proteins as well as the egg ES and soluble egg antigen (SEA) proteins were characterised by LC-MS/MS. To limit the number of false identifications we appended 116 sequences of common contaminants to the *S. haematobium* protein database (Bioproject PRJNA78265, www.parasite.wormbase.org/), and only proteins identified with ≥ 2 peptides were considered as positively identified. The extract with the highest number of identified proteins was the adult tegument (662, [S5 Table](#)), followed by adult ES (239, [S6 Table](#)), SEA (210, [S7 Table](#)) and egg ES (138, [S8 Table](#)). The adult tegument sample also had the highest number of unique (only present in this extract) proteins (430), followed by SEA (61), adult ES (48) and egg ES (19) ([Fig 1A](#)). Furthermore, adult tegument and adult ES proteins had 85 proteins in common, while only 57 proteins were commonly identified in all datasets ([Fig 1A](#)).

The top 25 most abundant proteins in each extract selected based on the spectral count provided a similar profile: 12 proteins were uniquely identified in adult tegument extract, followed by SEA (11) and adult ES and egg ES (8 each) ([Fig 1B](#)). Only 3 proteins were commonly identified in all datasets (MS3_04307.1, fatty acid binding protein—FABP; MS3_11229.1, triose phosphate isomerase; and MS3_06482.1, glutathione S-transferase—GST). Unique proteins

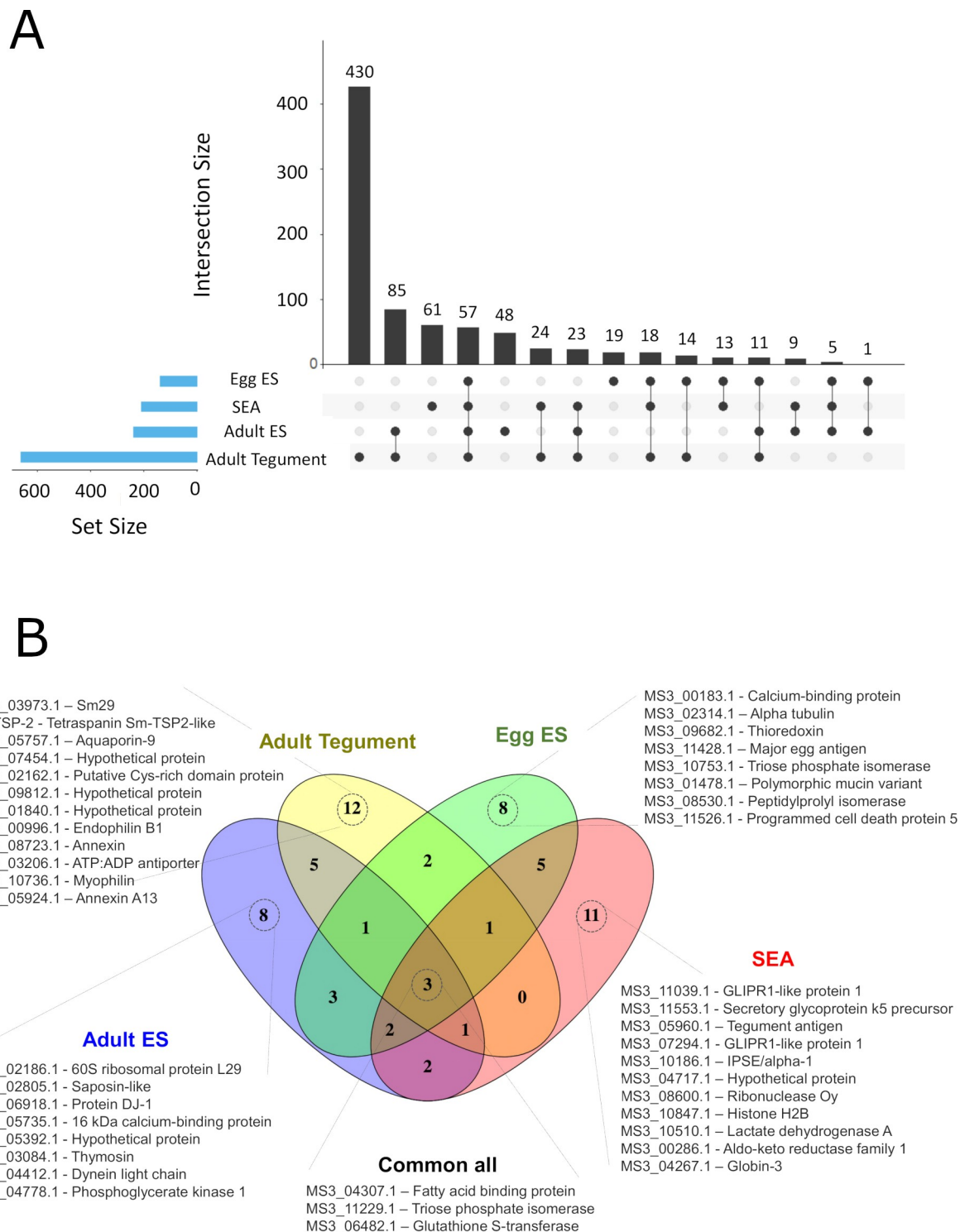


Fig 1. Proteins identified in the different *Schistosoma haematobium* proteomes. (A) The numbers and intersections of the identified proteins from different *Schistosoma haematobium* proteomes was visualised using an Upset plot. Connected dots display shared proteins between datasets, and the total number of proteins identified in a particular dataset is indicated in the set size. (B) Venn diagram representing the intersection of the top 25 most abundant proteins from each dataset based on spectrum counting.

<https://doi.org/10.1371/journal.pntd.0007362.g001>

identified from the adult tegument included three hypothetical proteins (MS3_07454.1, MS3_09812.1 and MS3_01840.1), one tetraspanin (*Sh*-TSP-2) and two annexins (MS3_08723.1 and MS3_05924.1), among others (Fig 1B). Among the proteins uniquely identified in SEA, we found two proteins of the glioma pathogenesis related-1 (GLIPR1) subfamily (MS3_11039.1 and MS3_07294.1), IPSE/alpha-1 protein (MS3_10186.1), as well as several enzymes (MS3_08600.1, MS3_10510.1 and MS3_00286.1) (Fig 1B). A calcium binding protein (MS3_00183.1), alpha-tubulin (MS3_02314.1), thioredoxin (MS3_09682.1) and a major egg antigen (MS3_11428.1) were the most highly represented proteins uniquely identified in egg ES, while uniquely identified proteins in adult ES included saposin (MS3_02805.1) and thymosin (MS3_03084.1).

Protein families and functions in the different proteomes of *Schistosoma haematobium*

A Pfam analysis was performed on the different proteomes from *S. haematobium* and the top 25 most represented protein families for each dataset were visualized in a heatmap (Fig 2). Five different protein domains (three different EF-hand-like domains, a tetratricopeptide repeat and an AAA domain) were highly represented in all proteomes, while three domains were exclusively found in proteins identified in the tegument of *S. haematobium* (ADP-ribosylation, Ras family and Ras of Complex Roc) (Fig 2). Similarly, one domain was found in proteins exclusively identified in the egg ES (KH domain), while other cytoskeletal and redox domains (e.g. redoxin, tropomyosin and glutathione S-transferase) were more abundant in proteins from egg ES and SEA. Interestingly, two immunoglobulin domains were highly represented in the proteins identified from the adult ES and also found in the adult tegument but not present in any of the egg proteomes (Fig 2).

The GO analysis was performed using Blast2GO [30] and biological processes and molecular functions plotted using ReviGO [31]. Several metabolic processes were highly represented in adult ES (S9 Table), together with “gluconeogenesis” and “proteolysis” (Fig 3A). An “oxidation-reduction process” was common between adult ES and tegument proteomes, while the tegument proteins were also involved in “phosphorylation” (Fig 3C, S10 Table). Regarding molecular functions inferred from adult ES and tegument proteins, both datasets were enriched in “oxidoreductase activity”, “ATP binding”, “transferase activity” and “cytoskeletal protein binding”, although presenting different scores and frequencies (Fig 3C and 3D).

Both egg ES and SEA had distinct profiles with regards to the terms associated with biological process (S11 and S12 Tables, respectively). Despite both proteomes contain proteins involved in “gluconeogenesis” and “proteolysis”, they had different associated metabolic processes (Fig 4A and 4C). Two molecular function terms were similar among all proteomes (“oxidoreductase activity”, “ATP binding”), while “protein binding” was exclusively found in egg ES and SEA (Fig 4B and 4D). In addition, the egg ES proteome was enriched in proteins involved in “kinase activity” and “calcium ion binding”, while the molecular functions in SEA proteins related to “transferase activity” and “metal ion binding” (Fig 4B and 4D).

Schistosoma haematobium secreted and tegument proteins have potential as diagnostic markers of infection

Using the different *S. haematobium* proteomes characterized in this study, we analyzed antibody responses in the urine of *S. haematobium*-infected individuals (Fig 5A–5D) to the various parasite extracts. Furthermore, the AUC generated from the ROC curves were used to determine the sensitivity and specificity of each antibody response, and the predictive value of infection (Fig 6A–6D). Antibodies to all extracts were significantly reactive in the urine of all

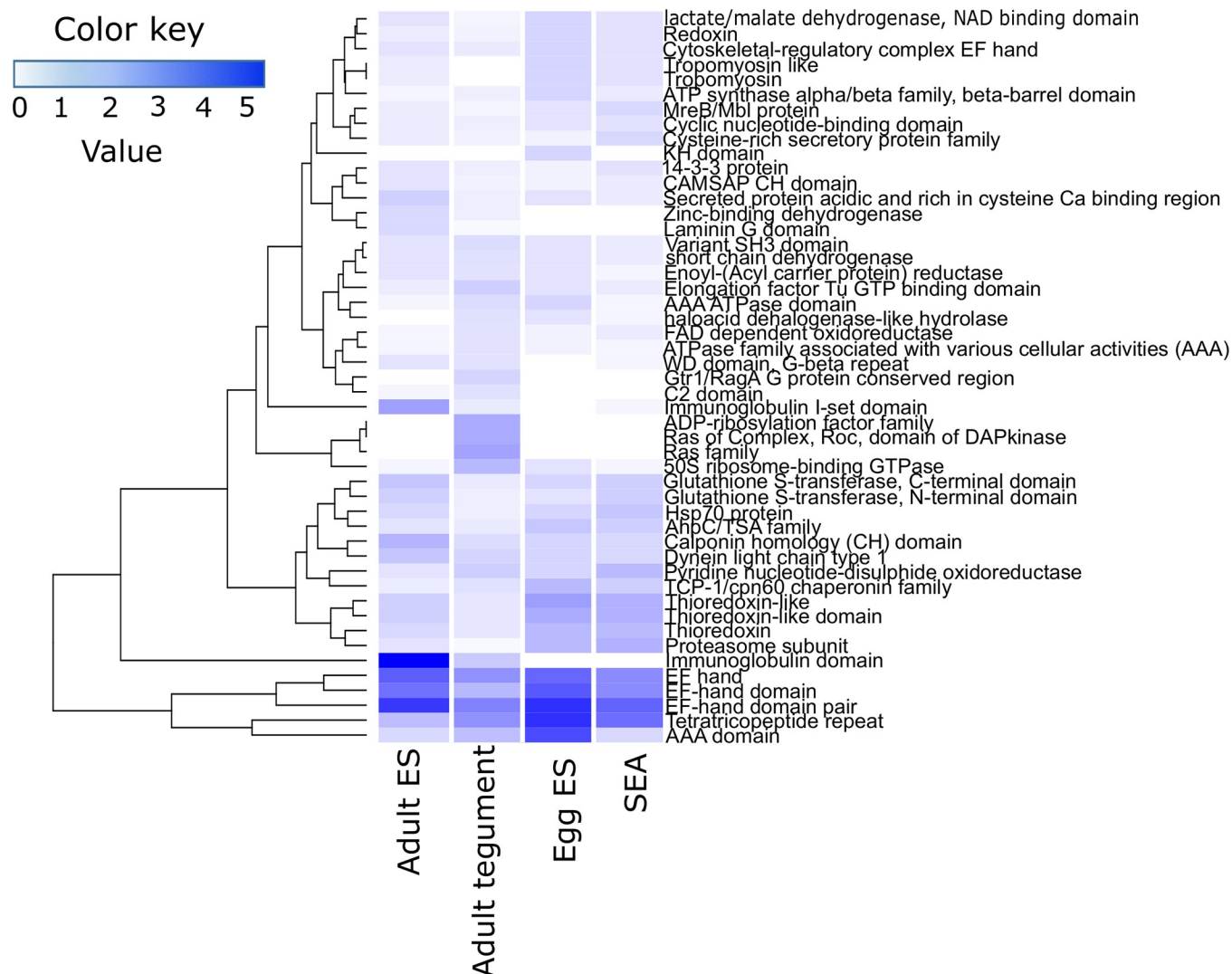


Fig 2. Pfam analysis. The top 25 most represented protein families from every proteome analysed were visualised using a heatmap. Values represent the abundance (in percentage) of each protein family relative to the total number of protein families present in each proteome.

<https://doi.org/10.1371/journal.pntd.0007362.g002>

subjects with a high egg count (>50 eggs/ml) compared to non-endemic negative controls with the most significant reactivity being to adult fluke ES and SEA ($p < 0.0001$), although adult tegument and egg ES samples were also highly reactive ($p < 0.001$). Antibodies in the urine of subjects with a medium intensity infection (11–50 eggs/ml) significantly reacted to all extracts, compared to non-endemic negative controls, with the most significant reactivity being to adult ES and tegument extracts ($p < 0.01$). Interestingly, the antibody levels in the urine of subjects with low intensity infection (0.3–10 eggs/ml) were significantly elevated only for adult ES ($p < 0.05$) and not for other extracts tested, while subjects with a very low level of infection (egg-negative by microscopy but CAA-positive) had significantly elevated antibody levels to only adult tegument and SEA extracts ($p < 0.05$).

The highest predictive value of infection in subjects with a high egg count was generated by the antibody response to adult ES (AUC 0.9954), followed by adult tegument (0.977), egg ES (0.9288) and SEA (0.9101) antibody responses. In the case of subjects with a medium egg count, the highest predictive value of infection was also observed with adult ES antibody

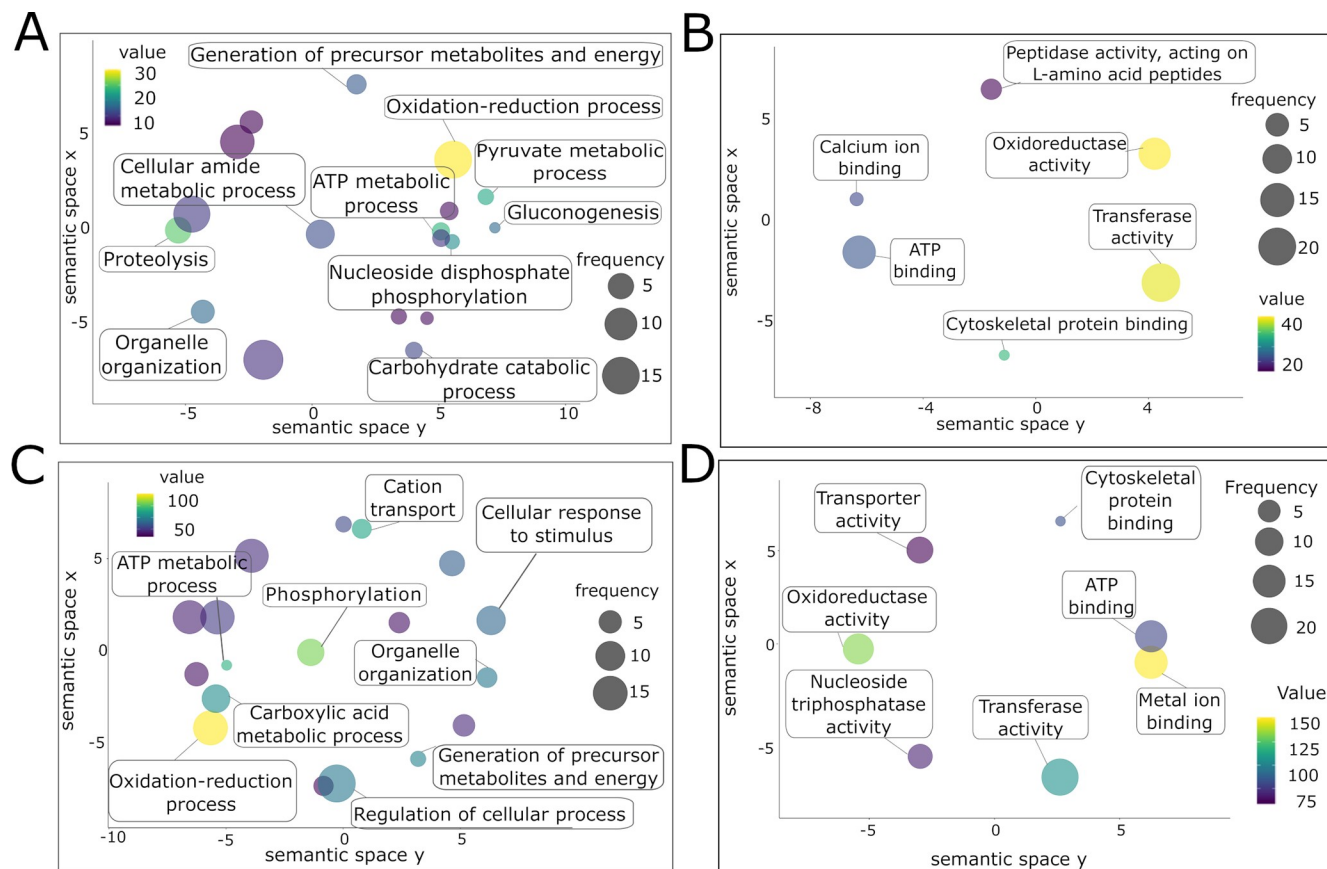


Fig 3. Gene ontology analysis of the adult excretory/secretory (ES) and tegument protein preparations from *Schistosoma haematobium*. Gene ontologies of proteins from *S. haematobium* adult ES (A, B) and tegument (C, D) were ranked by nodescore (Blast2GO) and plotted using ReviGO. Figure shows Biological Processes (A, C) and Molecular Function (B, D). Semantically similar GO terms plot close together, circle size denotes the frequency of the GO term from the underlying database, and increasing heatmap score signifies increasing nodescore from Blast2GO.

<https://doi.org/10.1371/journal.pntd.0007362.g003>

responses (0.9707), followed by adult tegument (0.8661), SEA (0.7691) and egg ES (0.7449). This was again the case for subjects with a low egg count; the predictive value of infection for adult ES, adult tegument, SEA and egg ES antibody responses being 0.7612, 0.6518, 0.702 and 0.5435, respectively. Interestingly, the highest predictive value of infection in microscopy egg negative but CAA positive subjects was generated with the antibody response to adult tegument extract (0.9762), followed by SEA (0.869), adult ES (0.8333) and egg ES (0.5357).

Discussion

Despite accounting for almost two-thirds of all cases of schistosomiasis [32], *S. haematobium* remains the most neglected schistosome of medical importance in terms of laboratory research effort, partly due to the lack of proteomic and genomic information (which underpin knowledge of the parasite's biological and pathogenic processes) [8]. Several versions of the *S. mansoni* genome have been published [33, 34], as well as numerous high-throughput proteomes from different life stages [22, 24–26, 35, 36]; however, only one version of the genome and very few proteins have been characterised for the causative agent of urinary schistosomiasis [37]. Indeed, a thorough examination of the *S. haematobium* genome failed to detect the presence of a homologue of *Sm*-TSP-2 in this parasite, despite this protein being found in the proteomic datasets described herein, its abundance in *S. mansoni* proteomes [38] and the

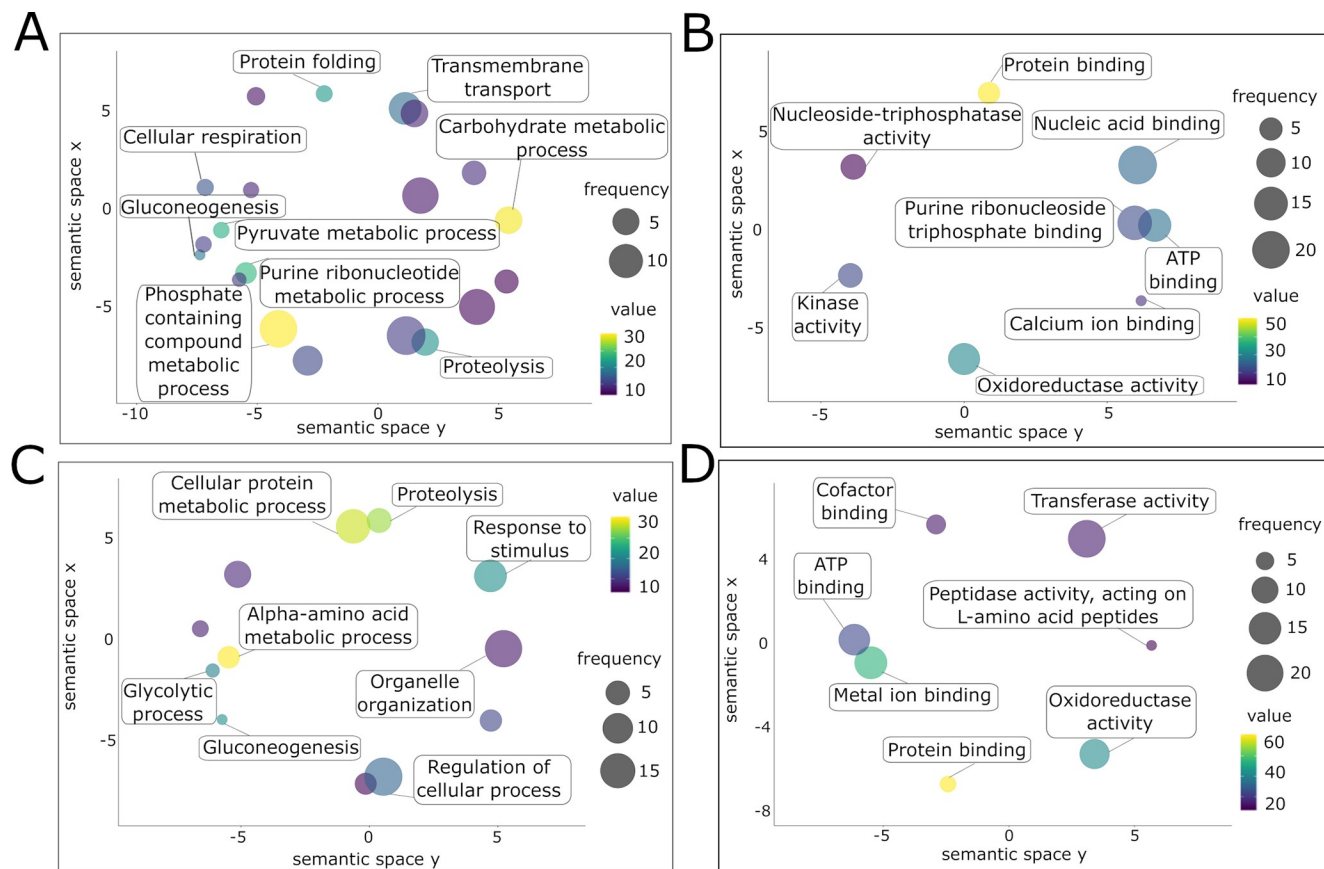


Fig 4. Gene ontology analysis of the egg excretory/secretory (ES) and soluble egg antigens (SEA) from *Schistosoma haematobium*. Gene ontologies of proteins from *S. haematobium* egg ES (A, B) and SEA (C, D) were ranked by nodescore (Blast2GO) and plotted using ReviGO. Figure shows Biological Processes (A, C) and Molecular Function (B, D). Semantically similar GO terms plot close together, circle size denotes the frequency of the GO term from the underlying database, and increasing heatmap score signifies increasing nodescore from Blast2GO.

<https://doi.org/10.1371/journal.pntd.0007362.g004>

presence of numerous homologs in *S. japonicum* [39]. *Sm*-TSP-2 has proven to be one of the most effective vaccine candidates against schistosomiasis [40, 41] and has successfully undergone a phase I clinical trial [42]; confirmation of the existence of a *S. haematobium* orthologue is important, therefore, to determine whether a similar (or existing) vaccination strategy might be effective against this species of schistosome. This example highlights the importance of the availability of proteomic data to complement and, as is being increasingly reported, re-annotate existing genomes to provide a comprehensive picture of the makeup of an organism [43].

Proteins at the interface between the host and the parasite are believed to play a key role in host immune system modulation and parasite survival [43] and so characterisation of these molecules is desirable to further our knowledge of parasite biology and pathogenesis and intervention targets. For example, the proteins secreted and located on/in the tegument of the blood-dwelling adult fluke are immune-accessible, and egg ES molecules are involved in the generation of a tumorigenic environment [13] and can be studied to gain insight into host pathogenesis.

A total of 57 proteins were commonly identified in all proteomic datasets generated. Among these, MS3_04307.1 (FABP), MS3_11229.1 (triose phosphate isomerase) and MS3_06482 (GST) were highly abundant. All proteomes were enriched in three main protein families: AAA domains, EF-hand like domains and tetratricopeptide repeat (TPR) domains.

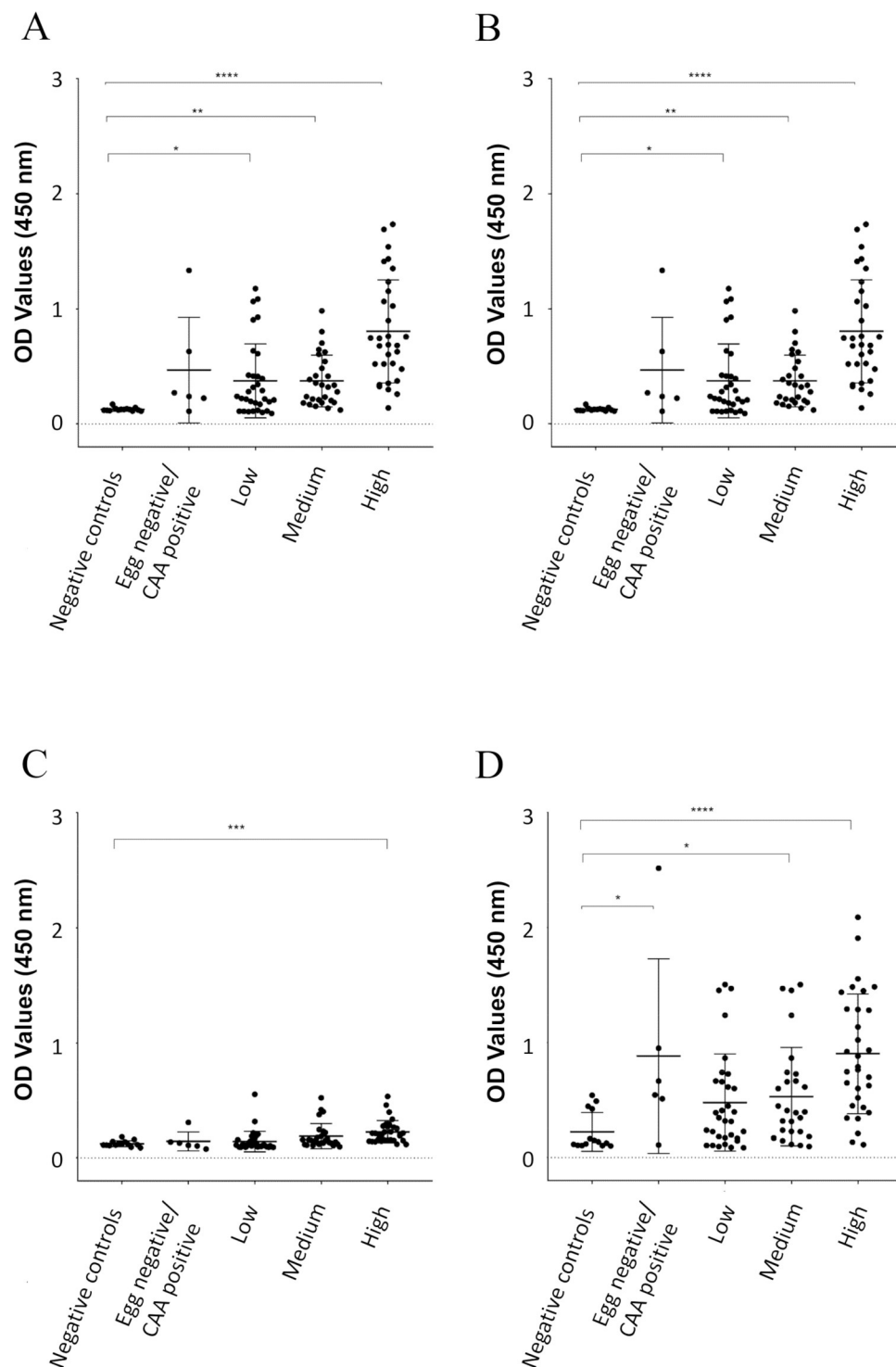


Fig 5. Urine antibody levels against different *Schistosoma haematobium* antigen preparations. OD values determined by ELISA using urine from individuals with different infection intensities based on egg counts (high, medium, low and egg negative/CAA positive) from an endemic area of Zimbabwe against different proteomes from *S. haematobium*. A control group of urine from non-endemic individuals who had never been infected with *S. haematobium* was included. Adult excretory/secretory (ES) products (A), tegument proteins (B), egg ES (C) and soluble egg antigen (D). Statistical analysis was performed using a non-parametric Kruskal-Wallis test with multiple comparisons by Dunn's post-test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

<https://doi.org/10.1371/journal.pntd.0007362.g005>

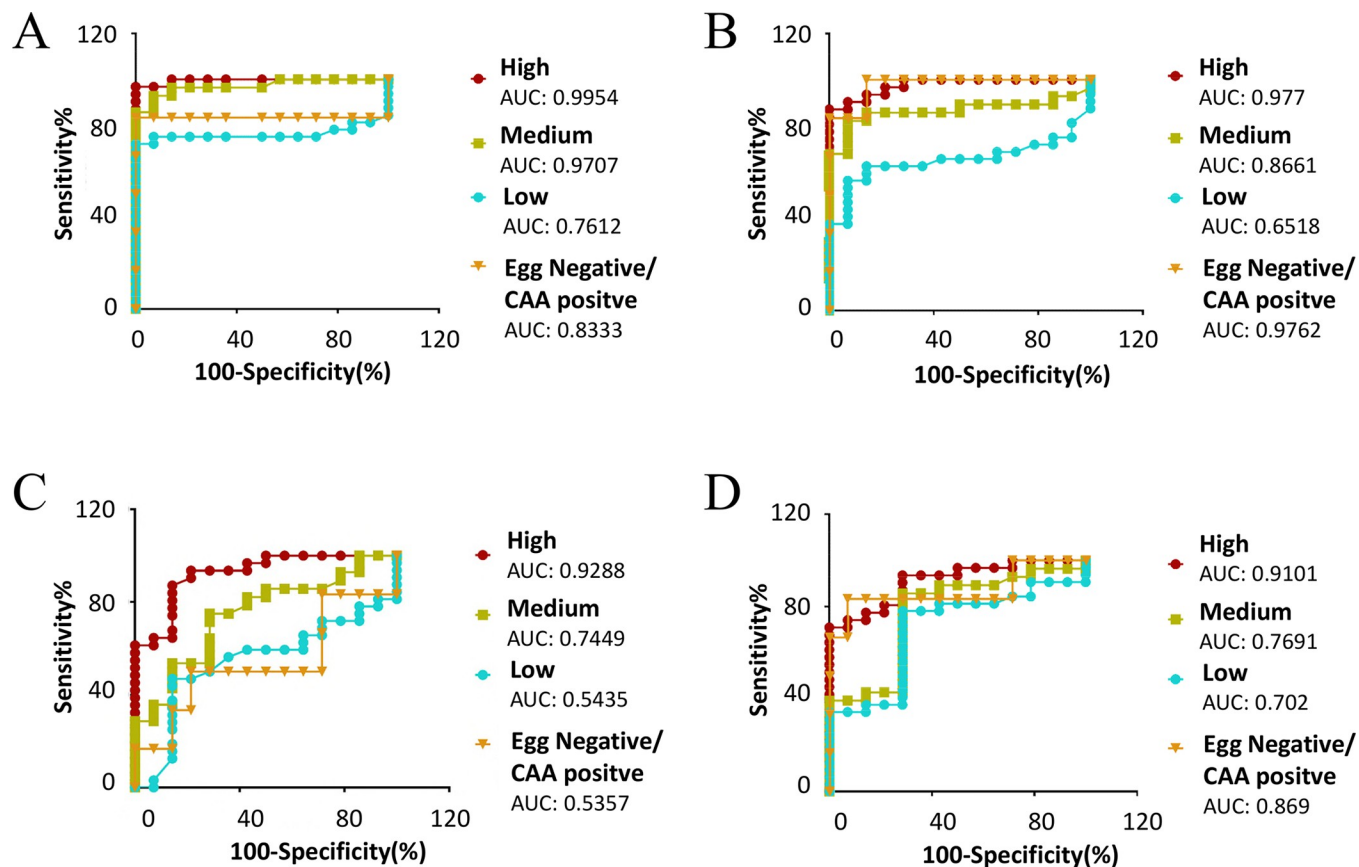


Fig 6. Receiver operating characteristic (ROC) curves. Areas under the ROC curve determine the predictive performance of adult excretory/secretory (ES) products (A), tegument proteins (B), egg ES products (C) and soluble egg antigen (SEA) (D) to detect antibodies in the urine of individuals with differing intensities (high, medium, low and egg negative/CAA positive) of *Schistosoma haematobium* infection. A control group of urine from non-endemic individuals who had never been infected with *S. haematobium* was included.

<https://doi.org/10.1371/journal.pntd.0007362.g006>

The high abundance of proteins containing these domains suggests a vital role for these motifs in parasite development and/or survival. While AAA domain-containing proteins such as dyneins participate in a variety of functions and are ubiquitously present in many organisms [44], EF-hand domain proteins are expressed almost uniquely by parasitic worms (highly abundant in trematodes), and have been suggested to be an effective drug target [45]. TPR is a structural repeat found in diverse proteins. In *Schistosoma mekongi* it is present within an O-glycosyltransferase that is abundantly expressed by female flukes [46], while in *S. japonicum*, the protein SJCHGC06661 was the highest immunoreactive protein in a protein array screened against the serum from *S. japonicum* infected patients. This protein contains six TPR domains, interacts with the heat shock protein (HSP) complex and has been suggested as a potential drug or vaccine target [47].

The most complex mixture of proteins was identified in the adult tegument sample (662 proteins). Previous studies on *S. mansoni* schistosomula and *S. japonicum* adult flukes identified 450 and 360 proteins, respectively [22, 48], highlighting the complex composition of this structure, probably due to an adaptive response to living in the harsh environment of the host circulatory system. The molecular functions associated with these proteins correspond to the functions observed in other studies [48]. For instance, proteins involved in several binding functions (the highest represented function in the proteins identified in the *S. japonicum* adult

tegument [48]) such as metal ion binding, cytoskeletal protein binding and ATP binding were highly represented in this study. Several molecular functions associated with protection against oxygen-reactive species and redox systems such as oxidoreductase activity and transferase activity were also highly represented, as expected in the tegument of a parasitic helminth which is under constant immune threat [49]. Among the most abundant proteins present in the tegument of *S. haematobium* adult flukes, we identified a putative GPI-anchored protein similar to Sm29 (MS3_03973.1), a tetraspanin similar to *Sm*-TSP-2 (*Sh*-TSP-2), an aquaporin (MS3_05757.1), as well as two different annexins (MS3_08723.1 and MS3_05924.1) and several hypothetical proteins (MS3_07454.1, MS3_09812.1 and MS3_01840.1). Sm29 has been shown to be an immunoregulatory molecule able to control inflammatory mucosal diseases with both Th1 and Th2 immune response profiles [50] and can interact with CD59, an inhibitor of the membrane attack complex (MAC), which could contribute to immune evasion [51]. Tetraspanins have a role in maintaining the structure of the tegument and providing stability [52]. Interestingly, Sm29 and *Sm*-TSP-2, as well as the above-mentioned GST and FABP, are important vaccine candidates [27]. Characterising the rest of the identified proteins, particularly the hypothetical proteins, will improve the rational selection of vaccine candidates for *S. haematobium* infection.

Using an older version of the *S. haematobium* proteome present in WormbaseParasite, only 379 proteins were predicted to be secreted by *S. haematobium*, making it the parasitic helminth with the lowest number of predicted secreted proteins of the 44 species assessed [53]. In our analysis, we found 239 proteins in the adult ES products, which accounts for more than 63% of the total predicted secretome. The most abundant domains identified by Cuesta-Astro in the predicted secretome were PF00053 (Laminin-EGF), which was not highly represented in our analysis, and PF13895 (Immunoglobulin domain), which was one of the most abundant families in the adult ES dataset. We also identified the Immunoglobulin I-set domain as an abundant protein in the adult ES products, which is also one of the most commonly occurring proteins in the predicted helminth secretomes [53], and is present in cell adhesion molecules. Interestingly, some of the identified proteins such as saposin, dynein and tetraspanins have been previously identified in the extracellular vesicles (EVs) of *S. mansoni* [54, 55]. EVs have not yet been described from *S. haematobium*, although it is tempting to speculate that some of the proteins we identified in ES products are of vesicular origin.

Egg secretions from trematodes likely result from an active secretion process of the miracidium enclosed within the egg and, in the case of *S. haematobium*, the proteins secreted by the egg might not only reflect the biology of the miracidium stage, but might also be of importance in the development of fibrosis and cancer. Among the top 25 most abundant proteins in all datasets, 8 proteins were uniquely identified in the egg secretions, including calcium-binding protein (MS3_00183.1), alpha-tubulin (MS3_02314.1), thioredoxin (MS3_09682.1), polymorphic mucin variant (MS3_01478.1) and the major egg antigen (MS3_11428.1). The calcium-binding protein is orthologous (90% identity) to *S. mansoni* calmodulin, which has been reported as an essential protein for egg development [56]. Indeed, calmodulin inhibitors can disrupt egg hatching and interrupt miracidium transformation into sporocyst [57, 58]. The polymorphic mucin variant has been identified in the miracidium of *S. mansoni* as having a key role in invertebrate host/parasite interactions [59], and a family of *S. mansoni* polymorphic mucins has been hypothesized to act as a smoke-screen blocking pattern recognition receptors, thus avoiding recognition by the host immune system [59].

Different proteins with a role in defence against oxidative stress, such as GST and thioredoxin, were abundant in the egg ES proteome; as well as two triose phosphate isomerases, which are proteins implicated in glycolysis and energy production. Interestingly, two homologs of programmed cell death protein 5 (PCDC5; MS3_11526.1) and the major egg antigen

also known as Smp40 (MS3_11428.1) were also uniquely identified in the egg ES and are among the top 25 most abundant proteins. PCDC5 has a wide variety of biological functions, including programmed cell death and immune regulation [60], and a decrease in the levels of this protein has been associated with multiple types of cancers [60]. Smp40 is associated with reduced collagen deposition, decreased fibrosis, and granuloma formation inhibition [61]. Active secretion of these two molecules might reflect an effort from the parasite to minimize the potential fibrosis and cell proliferation induced by other proteins present in the egg shell.

Two GLIPR1 proteins (MS3_11039.1 and MS3_07294.1) were uniquely identified in the SEA dataset and were among the most abundant proteins based on spectrum counting. These proteins belong to the sperm-coating proteins/Tpx1/Ag5/PR-1/Sc7 (SCP/TAPS) family, and have been identified in other trematodes including *Fasciola hepatica* [62] as well as other helminths [63]. Although the exact function for these proteins is still unknown, this family of proteins is expanded in the genomes and secreted proteomes of clade IV and V nematodes [64–68] and is believed to play specific biological functions in the host including defence against host attack and determination of lifespan [69]. The presence of these proteins in the eggs of *S. haematobium* is intriguing and their roles in this life stage warrants exploration.

The IPSE/alpha-1 glycoprotein (MS3_10186.1) is a well characterized molecule from the eggs of *S. mansoni*. It is located in the subshell area of mature eggs [70] and is a potent driver of IL-4-mediated Th2 responses [70–73]. Furthermore, it induces a potent anti-inflammatory response by inducing regulatory B cells to produce IL-10 [71]. The *S. haematobium* homolog of IPSE/alpha-1 (H-IPSE) infiltrates host cells and translocate to the nucleus [19] and, interestingly, has been shown to alleviate the symptoms of chemotherapy-induced hemorrhagic cystitis [20].

With a view to providing the first steps towards the characterisation of novel molecules that could be used for immunodiagnosis of *S. haematobium* infection, we assessed the levels of IgG antibodies present in the urine of patients from an endemic area in Zimbabwe against the different extracts analysed in the study. Even though the presence of antibodies in serum or urine cannot be used to differentiate previous and current infections, the detection of immunoglobulins against parasites is still a useful diagnostic tool for surveillance, and can be helpful for evaluating the effectiveness of control programmes [74]. Antibody detection diagnostics can also be used to complement less sensitive point-of-care (POC) diagnostic tests, such as microscopic detection of eggs in urine, in areas of low transmission. Some efforts have been made to develop highly sensitive and specific tests for *S. haematobium* infection. For instance, diverse parasite proteomes such as SEA, cercarial antigen preparation (CAP) and soluble worm antigen preparation (SWAP) have been tested [75]. In this study, SEA and CAP were more reactive than SWAP, which is in accordance with previous findings [74, 76–78]. The fact that SWAP is not a good extract from a diagnostic point of view could be because it is highly abundant in intracellular proteins that will never be exposed to host antibodies. The detection of antibodies against SEA and SWAP in the urine of human subjects with schistosomiasis has been previously reported [74, 77, 78], but the presence and diagnostic utility of urine antibodies to adult fluke tegument and ES have not been described. Herein, we highlight the diagnostic capability of arguably more immunologically relevant proteomes, in addition to SEA, to detect antibodies in the urine of infected individuals, and show that the adult fluke ES and tegument preparations were more reactive than SEA for all of the cohorts studied. Antibodies can be detected in urine of subjects infected with both *S. mansoni* and *S. japonicum* (where eggs are passed in the stool instead of urine), so the presence of antibodies in urine of subjects infected with *S. haematobium* is likely due to clearance of IgG from circulation through the kidneys as opposed to vascular leakage due to bladder pathology caused by parasite eggs. However, despite the fact that the prevalence of *S. mansoni* compared to *S. haematobium* infection is very low in

Zimbabwe, especially in the areas where the samples were collected [75, 79], we cannot exclude a role for *S. mansoni* infection in the responses observed. Future studies aimed at characterising species-specific diagnostic molecules will address this point.

We have provided the first comprehensive high-throughput analysis of the proteins present in different *S. haematobium* proteomes of importance in host/parasite interactions, thereby facilitating a snapshot of the molecular biology of the parasite and how it interacts with the host. In addition, we have identified adult fluke ES and tegument extracts as best predictors of infection when probed with antibodies from the urine of infected human subjects. This combined study of proteomic characterisation and serodiagnostic analyses provide the first steps towards the characterisation of novel molecules that could serve as tools for the control and evaluation of *S. haematobium* infection in Sub-Saharan Africa.

Materials and methods

Ethics

The collection of urine from individuals from Zimbabwe was approved by the Medical Research Council of Zimbabwe; Approval MRCZ/A/1710.

All experimental procedures involving animals reported in the study were approved by the James Cook University (JCU) animal ethics committee (ethics approval number A2391). The study protocols were in accordance with the 2007 Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and the 2001 Queensland Animal Care and Protection Act.

Human *Schistosoma haematobium* parasitology

Urine samples from naturally infected individuals were collected on three consecutive days for parasitological examinations. *S. haematobium* infection was detected by microscopic examination of the parasite eggs in 10ml urine, processed using the standard urine filtration method [80]. For each participant, infection intensity was expressed as the arithmetic mean of egg counts per 10 mL urine of samples collected on consecutive days. Urines were stratified according to WHO classification as having either a high (>50 eggs/10 ml of urine), medium (11–49 eggs/10ml of urine) or low (0.3–10 eggs/10 ml of urine) level of infection [81]. Egg negative urines (0 eggs/10 ml of urine) were tested for the presence of circulating anodic antigen (CAA—a more sensitive diagnostic test than microscopic detection of eggs in urine) using the UCAA2000 (wet format) as described previously [82] to confirm the presence or absence of infection.

Parasite material

S. haematobium-infected *Bulinus truncatus* snails were provided by the National Institute of Allergy and Infectious Diseases (NIAID) Schistosomiasis Resource Center for distribution through BEI Resources, NIAID, National Institutes of Health (NIH), USA: *S. haematobium*, Egyptian strain, exposed *B. truncatus* subsp. *truncatus*, NR-21965. Snails were removed from their tank, washed with water and transferred to a petri dish without light or water at 27°C for 90 minutes. They were then rinsed again and water was added to an approx. depth of 3–5 mm. Snails were then placed under direct light maintaining a temperature of 28–30°C for 1–2 hours. The water was then transferred to a new petri dish through a sieve with 20 µm pore size to concentrate cercariae. Water was transferred from the snails to the sieve every 20 min three more times while cercariae continued to be shed. Mice (6 week-old Balb/c) were infected with 1,000 cercariae by tail penetration and adult worms were recovered by vascular perfusion at 16

weeks p.i. [83]. Eggs from livers of perfused mice were isolated according to the method of Dalton et al. [84] to obtain highly purified ova that were free of host debris.

Isolation of adult excretory/secretory products

Two hundred freshly perfused adult fluke pairs were washed 3× in serum-free Basch media supplemented with 4× antibiotic/antimycotic (10,000 units/mL of penicillin, 10,000 µg/mL of streptomycin, and 25 µg/mL of Amphotericin B) (AA) (Thermo Fisher Scientific, USA) [85] followed by incubation in the same medium at 37°C, 5% CO₂ at a density of ~50 fluke pairs in 4 mL of medium for 7 days. Culture medium was changed after 4 hours and discarded to minimize the presence of host contaminants from regurgitating flukes. Media was subsequently changed every day and dead flukes removed from the plate to avoid contamination with somatic proteins. The excretory/secretory (ES) material was collected each day, centrifuged at 500 g, 2,000 g and 4,000 g to remove parasite debris, buffer exchanged in PBS, concentrated using a 10 kDa spin concentrator (Merck Millipore, USA), protein quantified by BCA (Thermo Fisher Scientific, USA), aliquoted and stored at -80°C until use.

Tegument extraction

Extraction of tegument proteins from adult flukes was performed using the freeze/thaw/vortex technique as described previously [22, 86]. Briefly, two batches of 50 freshly perfused adult fluke pairs were washed 3× in PBS, frozen at -80°C, thawed on ice, washed in TBS (10 mM Tris/HCl, 0.84% NaCl, pH 7.4) and incubated for 5 min on ice in 10 mM Tris/HCl, pH 7.4. Each sample was vortexed for 5 × 1 s bursts, the tegument extract pelleted at 1,000 g for 30 min and solubilized three times in 200 µl of solubilizing solution (0.1% (w/v) SDS, 1.0% (v/v) Triton X-100 in 40 mM Tris, pH 7.4) with pelleting at 15,000 g between each wash. The washes were combined and buffer exchanged and concentrated as described above.

Egg excretory/secretory products and soluble egg antigen

Purified eggs were cultured in serum-free Basch medium supplemented with 4× AA (50,000 eggs/5 ml) for 72 hours at 37°C with 5% CO₂. Medium containing egg ES products was harvested every 24 hours and processed as described for adult ES. Approximately 400,000 eggs were used for ES generation. To obtain SEA, freshly isolated eggs in PBS (100,000/ml) were homogenized in a hand-held Potter-Elvehjem glass homogeniser (15 ml capacity), centrifuged at 200 g for 20 min at 4°C and then the supernatant centrifuged at 100,000 g for 90 min at 4°C [83]. The supernatant was then processed as described for adult ES.

OFFGEL electrophoresis

OFFGEL fractionation was performed for adult ES, adult tegument and egg ES samples as described by Sotillo et al. [64]. A total of 100 µg of protein was resuspended in 50 mM NH₄CO₃, 20 mM DTT and incubated at 65°C for 1 h. Alkylation was then achieved by adding iodoacetamide (IAM) to 55 mM and incubating the solution for 40 min in darkness at room temperature (RT). A final incubation with 100 mM DTT was performed at RT before adding 2 µg of trypsin and incubating overnight at 37°C. Fractionation of samples was performed using a 3100 OFFGEL fractionator and OFFGEL kit (pH 3–10; 24-well format) (Agilent Technologies, Australia) according to the manufacturer's protocols. Briefly, resultant peptides were diluted in peptide-focusing buffer, 150 µL of sample was loaded into each of the 24 wells and the sample was focused in a current of 50 µA until 50 kilovolt hours was reached. Each of the fractions was collected, dried under a vacuum centrifuge and resuspended in 10 µL of 0.1%

TFA before being desalted using a Zip-Tip (Merck Millipore, USA). Finally, samples were dried again under a vacuum centrifuge and stored at -80°C until use.

Mass spectrometry

Due to limited availability of material because of the difficulty in maintaining *S. haematobium* in mice, different numbers of biological replicates were run for each sample. A total of 48 offgel fractions from 2 different biological replicates (24 samples each) were run for adult ES, 48 offgel fractions from 2 different biological replicates (24 samples each) for adult tegument proteins, 24 offgel fractions (from 1 replicate) for egg ES and 2 samples (from 2 replicates) for SEA.

All samples were analyzed by LC-MS/MS on a Shimadzu Prominence Nano HPLC coupled to a Triple ToF 5600 mass spectrometer (ABSCIEX) using a nano electrospray ion source. Samples were resuspended in 30 µl of solvent A (0.1% formic acid (aq)) and fifteen µl was injected onto a 50 mm x 300 µm C18 trap column (Agilent Technologies) at a rate of 60 µl/min. The samples were desalted on the trap column for 6 minutes using solvent A at the same rate, and the trap column was then placed in-line with an analytical 150 mm x 100 µm 300SBC18, 3.5 µm nano HPLC column (Agilent Technologies) for mass spectrometry analysis. Peptides were eluted in a linear gradient of 2–40% solvent B (90/10 acetonitrile/ 0.1% formic acid (aq)) over 80 min at 500 nL/min flow rate after 6 minutes of wash at 2% solvent B and followed by a steeper gradient from 40% to 80% solvent B in 10 min. Solvent B was held at 80% for 5 min for washing the column and returned to 2% solvent B for equilibration prior to the next sample injection. Mass spectrometer settings were: ionspray voltage = 2,200V, declustering potential (DP) = 100V, curtain gas flow = 25, nebuliser gas 1 (GS1) = 12 and interface heater = 150°C. The mass spectrometer acquired 250 ms full scan TOF-MS data followed by 20 by 250 ms full scan product ion data in an Information Dependent Acquisition (IDA) mode. Full scan TOF-MS data was acquired over the mass range 300–1600 and for product ion ms/ms 80–1600. Ions observed in the TOF-MS scan exceeding a threshold of 150 counts and a charge state of +2 to +5 were set to trigger the acquisition of product ion, ms/ms spectra of the resultant 20 most intense ions. The data was acquired using Analyst TF 1.6.1 (ABSCIEX).

Database search and protein identification

Peak lists obtained from MS/MS spectra were identified using a combination of five search engines in SearchGUI version v3.3.3 [87] since it has been shown that combining multiple search engines increases the confidence of identified peptide spectrum matches (PSMs), distinct peptide sequences and proteins [88]. The search engines used were: X!Tandem version X!Tandem Vengeance (2015.12.15.2) [89], MS-GF+ version Release (v2018.04.09) [90], Comet version 2018.01 rev. 0 [91], MyriMatch version 2.2.140 [92] and Tide [93].

Protein identification was conducted against a concatenated target/decoy version of the *S. haematobium* proteome downloaded from Parasite Wormbase (version of 2017-05-WormBase - www.parasite.wormbase.org, 11,140 sequences) and concatenated to the common repository of adventitious proteins (cRAP, <https://www.thegpm.org/crap/>, 116 sequences) and the sequence of an antigen ortholog to *S. mansoni* TSP-2 that was obtained from GenBank (MK238557) (total of 11,257 (target) sequences). The decoy sequences were created by reversing the target sequences in SearchGUI. The identification settings were as follows: Trypsin specific with a maximum of 2 missed cleavages, 10.0 ppm as MS1 and 0.2 Da as MS2 tolerances; fixed modifications: Carbamidomethylation of C (+57.021464 Da), variable modifications: Deamidation of N (+0.984016 Da), Deamidation of Q (+0.984016 Da), Oxidation of M (+15.994915 Da), fixed modifications during refinement procedure: Carbamidomethylation of C

(+57.021464 Da), variable modifications during refinement procedure: Acetylation of protein N-term (+42.010565 Da), Pyrolidone from E (-18.010565 Da), Pyrolidone from Q (-17.026549 Da), Pyrolidone from carbamidomethylated C (-17.026549 Da). All algorithm specific settings are listed in the Certificate of Analysis available in [S1–S4](#) Tables.

Peptides and proteins were inferred from the spectrum identification results using Peptide-Shaker version 1.16.27 [94]. PSMs, peptides and proteins were validated at a 1.0% False Discovery Rate (FDR) estimated using the decoy hit distribution. All validation thresholds are listed in the Certificate of Analysis available in the [S1–S4](#) Tables. Mass spectrometry data along with the identification results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository [95] with the dataset identifiers PXD011137 and 10.6019/PXD011137.

Bioinformatic analysis of proteomic sequence data

The programs Blast2GO v5.2 [30] and HMMER v3.1b1 [96] were used to classify the proteins according to GO categories and Pfam domains respectively. Pfam domains were detected at the $P < 0.01$ threshold for the HMMER software. ReviGO was used to visualise GO terms using semantic similarity-based scatterplots [31]. The UpSetR package (v. 1.3.3) [97] was used to visualise the intersections of proteins between samples by producing UpSetR plots in R.

Enzyme-linked immunosorbent assay with human urine

The urine of infected individuals ($n = 98$) from an area in Zimbabwe endemic for *S. haematobium* infection was analyzed by ELISA using all *S. haematobium* protein preparations described earlier. Urine from Australian volunteer donors that had never travelled to schistosomiasis endemic areas was used as a negative control ($n = 14$).

Polystyrene microtiter plates (Greiner Bio-One, Austria) were coated overnight at 4°C with 50 µl/well of a 5 µg/ml solution of *S. haematobium* adult ES, adult tegument, SEA or egg ES in 0.1 M carbonate coating buffer, pH 9.6. The plates were washed three times with PBS/0.05% Tween20 (PBST) and blocked for two hours at 37°C using 5% skimmed milk in PBST, followed by three wash steps in PBST for 15 min each. Plates were then incubated with 50 µl of urine (diluted 1:5 in PBS) and incubated at 37°C for 1.5 h followed by 3 washes using PBST. Fifty µl of HRP-conjugated polyclonal anti-human IgG (Sigma-Aldrich) was added at dilution of 1:5,000 and incubated for 1 hour at 37°C. Finally, plates were washed 3× with PBST and incubated with 3,3',5,5'-tetramethylbenzidine (TMB, Thermo Fisher Scientific, USA) for 10 min at RT in the dark. The reaction was stopped by adding 3 M HCl and absorbance read at 450 nm using a POLARstar Omega spectrophotometer (BMG Labtech, Australia).

Statistical analysis

GraphPad Prism 7.0 was used for statistical analyses. Differences in antibody titers were analysed using the non-parametric Kruskal-Wallis test with Dunn's multiple comparisons test. Receiver Operating Characteristic (ROC) curves were used to calculate sensitivity, specificity and the area under the curve (AUC). The AUC is a global measure of diagnostic accuracy and indicates the probability of accurately identifying true positives, where a value of 0 indicates a perfectly inaccurate test and a value of 1 reflects a perfectly accurate test [98].

Supporting information

S1 Table. Certificate of analysis provided by PeptideShaker for the analysis of the *Schistosoma haematobium* adult secreted proteins.

(XLS)

S2 Table. Certificate of analysis provided by PeptideShaker for the analysis of the *Schistosoma haematobium* adult tegumental proteins.

(XLS)

S3 Table. Certificate of analysis provided by PeptideShaker for the analysis of the *Schistosoma haematobium* egg secreted proteins.

(XLS)

S4 Table. Certificate of analysis provided by PeptideShaker for the analysis of the *Schistosoma haematobium* soluble egg antigen proteins.

(XLS)

S5 Table. *Schistosoma haematobium* adult secreted proteins identified using SearchGUI and PeptideShaker.

(XLSX)

S6 Table. *Schistosoma haematobium* adult tegumental proteins identified using SearchGUI and PeptideShaker.

(XLSX)

S7 Table. *Schistosoma haematobium* egg secreted proteins identified using SearchGUI and PeptideShaker.

(XLSX)

S8 Table. *Schistosoma haematobium* soluble egg antigen proteins identified using SearchGUI and PeptideShaker.

(XLSX)

S9 Table. Annotation of the *Schistosoma haematobium* adult secreted proteins provided by Blast2GO.

(XLSX)

S10 Table. Annotation of the *Schistosoma haematobium* adult tegumental proteins provided by Blast2GO.

(XLSX)

S11 Table. Annotation of the *Schistosoma haematobium* egg secreted proteins provided by Blast2GO.

(XLSX)

S12 Table. Annotation of the *Schistosoma haematobium* soluble egg antigen proteins provided by Blast2GO.

(XLSX)

Acknowledgments

The authors thank the study participants, as well as the parents/legal guardians and their teachers in Zimbabwe for their support of this study. They are very grateful for the cooperation of the Ministry of Health and Child Welfare in Zimbabwe. For their technical support, they

would like to thank the members of the Department of Biochemistry at the University of Zimbabwe.

Author Contributions

Conceptualization: Javier Sotillo, Mark S. Pearson, Alex Loukas.

Formal analysis: Javier Sotillo.

Investigation: Javier Sotillo, Mark S. Pearson, Abena S. Amoah, Govert van Dam, Paul L. A. M. Corstjens, Francisca Mutapi.

Resources: Javier Sotillo, Luke Becker, Gebeyaw G. Mekonnen, Janice Murray, Takafira Mdu-luza, Francisca Mutapi.

Writing – original draft: Javier Sotillo, Mark S. Pearson.

Writing – review & editing: Javier Sotillo, Mark S. Pearson, Francisca Mutapi, Alex Loukas.

References

- McManus DP, Dunne DW, Sacko M, Utzinger J, Vennervald BJ, Zhou XN. Schistosomiasis. Nat Rev Dis Primers. 2018; 4(1):13. Epub 2018/08/11. <https://doi.org/10.1038/s41572-018-0013-8> PMID: 30093684.
- Colley DG, Bustinduy AL, Secor WE, King CH. Human schistosomiasis. Lancet. 2014; 383(9936):2253–64. [https://doi.org/10.1016/S0140-6736\(13\)61949-2](https://doi.org/10.1016/S0140-6736(13)61949-2) PMID: 24698483.
- Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. Lancet. 2006; 368(9541):1106–18. [https://doi.org/10.1016/S0140-6736\(06\)69440-3](https://doi.org/10.1016/S0140-6736(06)69440-3) PMID: 16997665.
- Hotez PJ, Alvarado M, Basanez MG, Bolliger I, Bourne R, Boussinesq M, et al. The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases. PLoS Negl Trop Dis. 2014; 8(7):e2865. Epub 2014/07/25. <https://doi.org/10.1371/journal.pntd.0002865> PMID: 25058013; PubMed Central PMCID: PMC4109880.
- Rollinson D, Knopp S, Levitz S, Stothard JR, Tchuente LA, Garba A, et al. Time to set the agenda for schistosomiasis elimination. Acta Trop. 2013; 128(2):423–40. Epub 2012/05/15. <https://doi.org/10.1016/j.actatropica.2012.04.013> PMID: 22580511.
- DALYs GBD, Collaborators H. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017; 390(10100):1260–344. Epub 2017/09/19. [https://doi.org/10.1016/S0140-6736\(17\)32130-X](https://doi.org/10.1016/S0140-6736(17)32130-X) PMID: 28919118; PubMed Central PMCID: PMC5605707.
- King CH. It's Time to Dispel the Myth of "Asymptomatic" Schistosomiasis. PLoS Negl Trop Dis. 2015; 9(2):e0003504. <https://doi.org/10.1371/journal.pntd.0003504> PMID: 25695740; PubMed Central PMCID: PMC4335065.
- Rinaldi G, Okatcha TI, Popratiloff A, Ayuk MA, Suttiprapa S, Mann VH, et al. Genetic Manipulation of *Schistosoma haematobium*, the Neglected Schistosome. PLoS Negl Trop Dis. 2011; 5(10):e1348. <https://doi.org/10.1371/journal.pntd.0001348> PMID: 22022628
- Lewis F. Schistosomiasis Current protocols in immunology. 2001;Chapter 19:Unit-19.1. <https://doi.org/10.1002/0471142735.im1901s28> PMID: 18432750.
- World Health Organization. Schistosomiasis (WHO, 2017).
- Boissier J, Grech-Angelini S, Webster BL, Allienne JF, Huyse T, Mas-Coma S, et al. Outbreak of urogenital schistosomiasis in Corsica (France): an epidemiological case study. Lancet Infect Dis. 2016; 16(8):971–9. Epub 2016/05/21. [https://doi.org/10.1016/S1473-3099\(16\)00175-4](https://doi.org/10.1016/S1473-3099(16)00175-4) PMID: 27197551.
- Kincaid-Smith J, Tracey A, de Carvalho Augusto R, Bulla I, Holroyd N, Rognon A, et al. Whole genome sequencing and morphological analysis of the human-infecting schistosome emerging in Europe reveals a complex admixture between *Schistosoma haematobium* and *Schistosoma bovis* parasites. bioRxiv. 2018.
- Ishida K, Hsieh MH. Understanding urogenital schistosomiasis-related bladder cancer: an update. Front Med. 2018; 5:223. Epub 2018/08/31. <https://doi.org/10.3389/fmed.2018.00223> PMID: 30159314; PubMed Central PMCID: PMC6104441.

14. Brindley PJ, Loukas A. Helminth infection-induced malignancy. *PLoS Pathog*. 2017; 13(7):e1006393. Epub 2017/07/28. <https://doi.org/10.1371/journal.ppat.1006393> PMID: 28750101; PubMed Central PMCID: PMC5531427.
15. Humans IWGotEoCRt. Biological agents. Volume 100 B. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012; 100(Pt B):1–441. Epub 2012/11/30. PMID: 23189750; PubMed Central PMCID: PMC54781184.
16. Barsoum RS. Urinary schistosomiasis: review. *J Adv Res*. 2013; 4(5):453–9. Epub 2013/09/01. <https://doi.org/10.1016/j.jare.2012.08.004> PMID: 25685452; PubMed Central PMCID: PMC54293885.
17. Chenault C, Hoang MP. An Unusual cervical finding. *Archives of Pathology & Laboratory Medicine*. 2006; 130(3):e37–e8. [https://doi.org/10.1043/1543-2165\(2006\)130\[e37:AUCF\]2.0.CO;2](https://doi.org/10.1043/1543-2165(2006)130[e37:AUCF]2.0.CO;2)
18. Michelle North MD, Irina Dubinchik MD, Adel Hamid MD, Salah Elderiny MD, Raja Sayegh MD. Association between cervical schistosomiasis and cervical cancer: A report of 2 cases. *J Reprod Med*. 2003; 48(12):4.
19. Pennington LF, Alouffi A, Mbanefo EC, Ray D, Heery DM, Jardetzky TS, et al. H-IPSE is a pathogen-secreted host nucleus infiltrating protein (infiltrin) expressed exclusively by the *Schistosoma haematobium* egg stage. *Infect Immun*. 2017. Epub 2017/09/20. <https://doi.org/10.1128/IAI.00301-17> PMID: 28923894; PubMed Central PMCID: PMC5695104.
20. Mbanefo EC, Le L, Pennington LF, Odegaard JI, Jardetzky TS, Alouffi A, et al. Therapeutic exploitation of IPSE, a urogenital parasite-derived host modulatory protein, for chemotherapy-induced hemorrhagic cystitis. *FASEB J*. 2018; 32(8):4408–19. Epub 2018/04/04. <https://doi.org/10.1096/fj.201701415R> PMID: 29613835; PubMed Central PMCID: PMC6044057.
21. Mutapi F. Helminth parasite proteomics: from experimental models to human infections. *Parasitology*. 2012; 139(9):1195–204. Epub 2012/03/28. <https://doi.org/10.1017/S0031182011002423> PMID: 22455721.
22. Sotillo J, Pearson M, Becker L, Mulvenna J, Loukas A. A quantitative proteomic analysis of the tegumental proteins from *Schistosoma mansoni* schistosomula reveals novel potential therapeutic targets. *Int J Parasitol*. 2015; 45(8):505–16. Epub 2015/04/26. <https://doi.org/10.1016/j.ijpara.2015.03.004> PMID: 25910674.
23. Perez-Sanchez R, Ramajo-Hernandez A, Ramajo-Martin V, Oleaga A. Proteomic analysis of the tegument and excretory-secretory products of adult *Schistosoma bovis* worms. *Proteomics*. 2006; 6 Suppl 1:S226–36. Epub 2006/03/03. <https://doi.org/10.1002/pmic.200500420> PMID: 16511809.
24. Braschi S, Borges WC, Wilson RA. Proteomic analysis of the schistosome tegument and its surface membranes. *Mem Inst Oswaldo Cruz*. 2006; 101 Suppl 1:205–12. PMID: 17308771.
25. Braschi S, Curwen RS, Ashton PD, Verjovski-Almeida S, Wilson A. The tegument surface membranes of the human blood parasite *Schistosoma mansoni*: a proteomic analysis after differential extraction. *Proteomics*. 2006; 6(5):1471–82. <https://doi.org/10.1002/pmic.200500368> PMID: 16447162.
26. Braschi S, Wilson RA. Proteins exposed at the adult schistosome surface revealed by biotinylation. *Mol Cell Proteomics*. 2006; 5(2):347–56. <https://doi.org/10.1074/mcp.M500287-MCP200> PMID: 16269422.
27. Tebeje BM, Harvie M, You H, Loukas A, McManus DP. Schistosomiasis vaccines: where do we stand? *Parasit Vectors*. 2016; 9(1):528. Epub 2016/10/08. <https://doi.org/10.1186/s13071-016-1799-4> PMID: 27716365; PubMed Central PMCID: PMC5045607.
28. ClinicalTrials.gov. Efficacy of vaccine sh28GST in association with praziquantel (PZQ) for prevention of clinical recurrences of *Schistosoma haematobium* pathology (Bilhvax). ClinicalTrials.gov Identifier: NCT00870649.
29. Riveau G, Deplanque D, Remoué F, Schacht A-M, Vodougnon H, Capron M, et al. Safety and immunogenicity of rsh28GST antigen in humans: Phase 1 randomized clinical study of a vaccine candidate against urinary schistosomiasis. *PLoS Negl Trop Dis*. 2012; 6(7):e1704. <https://doi.org/10.1371/journal.pntd.0001704> PMID: 22802974
30. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005; 21(18):3674–6. <https://doi.org/10.1093/bioinformatics/bti610> PMID: 16081474.
31. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLOS ONE*. 2011; 6(7):e21800. <https://doi.org/10.1371/journal.pone.0021800> PMID: 21789182
32. van der Werf MJ, de Vlas SJ, Brooker S, Looman CWN, Nagelkerke NJD, Habbema JDF, et al. Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. *Acta Tropica*. 2003; 86(2):125–39. [https://doi.org/10.1016/S0001-706X\(03\)00029-9](https://doi.org/10.1016/S0001-706X(03)00029-9).

33. Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. PLoS Negl Trop Dis. 2012; 6(1):e1455. <https://doi.org/10.1371/journal.pntd.0001455> PMID: 22253936
34. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, et al. The genome of the blood fluke *Schistosoma mansoni*. Nature. 2009; 460:352. <https://doi.org/10.1038/nature08160> PMID: 19606141
35. Cass CL, Johnson JR, Califf LL, Xu T, Hernandez HJ, Stadecker MJ, et al. Proteomic analysis of *Schistosoma mansoni* egg secretions. Mol Biochem Parasitol. 2007; 155(2):84–93. <https://doi.org/10.1016/j.molbiopara.2007.06.002> PubMed PMID: PMC2077830. PMID: 17644200
36. Figueiredo BC-P, Ricci ND, de Assis NRG, de Moraes SB, Fonseca CT, Oliveira SC. Kicking in the guts: *Schistosoma mansoni* digestive tract proteins are potential candidates for vaccine development. Front Immunol. 2015; 6:22. <https://doi.org/10.3389/fimmu.2015.00022> PubMed PMID: PMC4309203. PMID: 25674091
37. Young ND, Jex AR, Li B, Liu S, Yang L, Xiong Z, et al. Whole-genome sequence of *Schistosoma haematobium*. Nat Genet. 2012; 44:221. <https://doi.org/10.1038/ng.1065> PMID: 22246508
38. Castro-Borges W, Simpson DM, Dowle A, Curwen RS, Thomas-Oates J, Beynon RJ, et al. Abundance of tegument surface proteins in the human blood fluke *Schistosoma mansoni* determined by QconCAT proteomics. J Proteomics. 2011; 74(9):1519–33. <https://doi.org/10.1016/j.jprot.2011.06.011> PMID: 21704203
39. Cai P, Bu L, Wang J, Wang Z, Zhong X, Wang H. Molecular characterization of *Schistosoma japonicum* tegument protein tetraspanin-2: Sequence variation and possible implications for immune evasion. Biochem Biophys Res Commun. 2008; 372(1):197–202. <https://doi.org/10.1016/j.bbrc.2008.05.042> PMID: 18486598
40. Pearson MS, Pickering DA, McSorley HJ, Bethony JM, Tribolet L, Dougall AM, et al. Enhanced protective efficacy of a chimeric form of the schistosomiasis vaccine antigen Sm-TSP-2. PLoS Negl Trop Dis. 2012; 6(3):e1564. <https://doi.org/10.1371/journal.pntd.0001564> PMID: 22428079
41. Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, et al. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. Nat Med. 2006; 12:835. <https://doi.org/10.1038/nm1430> PMID: 16783371
42. Merrifield M, Hotez PJ, Beaumier CM, Gillespie P, Strych U, Hayward T, et al. Advancing a vaccine to prevent human schistosomiasis. Vaccine. 2016; 34(26):2988–91. <https://doi.org/10.1016/j.vaccine.2016.03.079> PMID: 27036511
43. Sotillo J, Toledo R, Mulvenna J, Loukas A. Exploiting helminth–host interactomes through big data. Trends Parasitol. 2017; 33(11):875–88. <https://doi.org/10.1016/j.pt.2017.06.011> PMID: 28734897
44. Erzberger JP, Berger JM. Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct. 2006; 35(1):93–114. <https://doi.org/10.1146/annurev.biophys.35.040405.101933> PMID: 16689629
45. Thomas Charlotte M, Timson David J. A mysterious family of calcium-binding proteins from parasitic worms. Biochem Soc Trans. 2016; 44(4):1005. <https://doi.org/10.1042/BST20150270> PMID: 27528745
46. Phuphisut O, Ajawatanawong P, Limpanont Y, Reamtong O, Nuamtanong S, Ampawong S, et al. Transcriptomic analysis of male and female *Schistosoma mekongi* adult worms. Parasit Vectors. 2018; 11(1):504. <https://doi.org/10.1186/s13071-018-3086-z> PMID: 30201055
47. Chen J-H, Zhang T, Ju C, Xu B, Lu Y, Mo X-J, et al. An integrated immunoproteomics and bioinformatics approach for the analysis of *Schistosoma japonicum* tegument proteins. J Proteomics. 2014; 98:289–99. <https://doi.org/10.1016/j.jprot.2014.01.010> PMID: 24448400
48. Zhang M, Hong Y, Han Y, Han H, Peng J, Qiu C, et al. Proteomic Analysis of tegument-exposed proteins of female and male *Schistosoma japonicum* worms. J Proteome Res. 2013; 12(11):5260–70. <https://doi.org/10.1021/pr400476a> PMID: 23909873
49. Zarowiecki M, Berriman M. What helminth genomes have taught us about parasite evolution. Parasitology. 2015; 142(Suppl 1):S85–S97. <https://doi.org/10.1017/S0031182014001449> PubMed PMID: PMC4413821. PMID: 25482650
50. Oliveira SC, Figueiredo BC, Cardoso LS, Carvalho EM. A double edged sword: *Schistosoma mansoni* Sm29 regulates both Th1 and Th2 responses in inflammatory mucosal diseases. Mucosal Immunol. 2016; 9:1366. <https://doi.org/10.1038/mi.2016.69> PMID: 27554296
51. White Bear J, Long T, Skinner D, McKerrow JH. Predictions of novel *Schistosoma mansoni*–human protein interactions consistent with experimental data. Sci Rep. 2018; 8(1):13092. <https://doi.org/10.1038/s41598-018-31272-1> PMID: 30166569

52. Tran MH, Freitas TC, Cooper L, Gaze S, Gatton ML, Jones MK, et al. Suppression of mrnas encoding tegument tetraspanins from *Schistosoma mansoni* results in impaired tegument turnover. PLoS Path. 2010; 6(4):e1000840. <https://doi.org/10.1371/journal.ppat.1000840> PMID: 20419145
53. Cuesta-Astroz Y, Oliveira FSd, Nahum LA, Oliveira G. Helminth secretomes reflect different lifestyles and parasitized hosts. International J Parasitol. 2017; 47(9):529–44. <https://doi.org/10.1016/j.ijpara.2017.01.007>.
54. Nowacki FC, Swain MT, Klychnikov OI, Niazi U, Ivens A, Quintana JF, et al. Protein and small non-coding RNA-enriched extracellular vesicles are released by the pathogenic blood fluke *Schistosoma mansoni*. J Extracell Vesicles. 2015; 4:28665. <https://doi.org/10.3402/jev.v4.28665> PMID: 26443722; PubMed Central PMCID: PMC4595467.
55. Sotillo J, Pearson M, Potriquet J, Becker L, Pickering D, Mulvenna J, et al. Extracellular vesicles secreted by *Schistosoma mansoni* contain protein vaccine candidates. Int J Parasitol. 2016; 46(1):1–5. <https://doi.org/10.1016/j.ijpara.2015.09.002> PMID: 26460238.
56. Taft AS, Yoshino TP. Cloning and functional characterization of two calmodulin genes during larval development in the parasitic flatworm *Schistosoma mansoni*. J Parasitol. 2011; 97(1):72–81. <https://doi.org/10.1645/GE-2586.1> PMID: 21348610
57. Katsumata T, Kohno S, Yamaguchi K, Hara K, Aoki Y. Hatching of *Schistosoma mansoni* eggs is a Ca²⁺/calmodulin-dependent process. Parasitol Res. 1989; 76(1):90–1. <https://doi.org/10.1007/BF00931079> PMID: 2622901
58. Kawamoto F, Shozawa A, Kumada N, Kojima K. Possible roles of cAMP and Ca²⁺ in the regulation of miracidial transformation in *Schistosoma mansoni*. Parasitol Res. 1989; 75(5):368–74. <https://doi.org/10.1007/BF00931132> PMID: 2542928
59. Roger E, Grunau C, Pierce RJ, Hirai H, Gourbal B, Galinier R, et al. Controlled chaos of polymorphic mucins in a metazoan parasite (*Schistosoma mansoni*) interacting with its invertebrate host (*Biomphalaria glabrata*). PLoS Negl Trop Dis. 2008; 2(11):e330. <https://doi.org/10.1371/journal.pntd.0000330> PMID: 19002242
60. Li G, Ma D, Chen Y. Cellular functions of programmed cell death 5. Biochim Biophys Acta Mol Cell Res. 2016; 1863(4):572–80. <https://doi.org/10.1016/j.bbamcr.2015.12.021>.
61. Abouel-Nour MF, Lotfy M, Attallah AM, Doughty BL. *Schistosoma mansoni* major egg antigen Smp40: molecular modeling and potential immunoreactivity for anti-pathology vaccine development. Mem Inst Oswaldo Cruz. 2006; 101:365–72.
62. Di Maggio LS, Tirloni L, Pinto AFM, Diedrich JK, Yates Iii JR, Benavides U, et al. Across intra-mammalian stages of the liver fluke *Fasciola hepatica*: a proteomic study. Sci Rep. 2016; 6:32796. <https://doi.org/10.1038/srep32796> PMID: 27600774
63. Cantacessi C, Hofmann A, Young ND, Broder U, Hall RS, Loukas A, et al. Insights into SCP/TAPS proteins of liver flukes based on large-scale bioinformatic analyses of sequence datasets. PLoS One. 2012; 7(2):e31164. <https://doi.org/10.1371/journal.pone.0031164> PMID: 22384000
64. Sotillo J, Sanchez-Flores A, Cantacessi C, Marcus Y, Pickering D, Bouchery T, et al. Secreted proteomes of different developmental stages of the gastrointestinal nematode *Nippostrongylus brasiliensis*. Mol Cell Proteomics. 2014; 13(10):2736–51. <https://doi.org/10.1074/mcp.M114.038950> PMID: 24994561
65. Tang YT, Gao X, Rosa BA, Abubucker S, Hallsworth-Pepin K, Martin J, et al. Genome of the human hookworm *Necator americanus*. Nat Genet. 2014; 46(3):261–9. <https://doi.org/10.1038/ng.2875> PMID: 24441737; PubMed Central PMCID: PMC3978129.
66. Morante T, Shepherd C, Constantinoiu C, Loukas A, Sotillo J. Revisiting the Ancylostoma Caninum Secretome Provides New Information on Hookworm–Host Interactions. Proteomics. 2017; 17(23–24):1700186. <https://doi.org/10.1002/pmic.201700186> PMID: 29052354
67. Hunt VL, Tsai IJ, Coghlan A, Reid AJ, Holroyd N, Foth BJ, et al. The genomic basis of parasitism in the Strongyloides clade of nematodes. Nat Genet. 2016; 48:299. <https://doi.org/10.1038/ng.3495> PMID: 26829753
68. Logan J, Manda SS, Choi Y-J, Field M, Eichenberger R, Mulvenna J, et al. Comprehensive analysis of human hookworm secreted proteins using a proteogenomic approach. bioRxiv. 2018.
69. Cantacessi C, Campbell BE, Visser A, Geldhof P, Nolan MJ, Nisbet AJ, et al. PMID: A portrait of the “SCP/TAPS” proteins of eukaryotes—Developing a framework for fundamental research and biotechnological outcomes. Biotech Adv. 2009; 27(4):376–88. <https://doi.org/10.1016/j.biotechadv.2009.02.005>.
70. Schramm G, Gronow A, Knobloch J, Wipperfsteg V, Grevelding CG, Galle J, et al. IPSE/alpha-1: A major immunogenic component secreted from *Schistosoma mansoni* eggs. Mol Biochem Parasitol. 2006; 147(1):9–19. <https://doi.org/10.1016/j.molbiopara.2006.01.003> PMID: 16480783

71. Haeberlein S, Obieglo K, Ozir-Fazalalikhan A, Chayé MAM, Veninga H, van der Vlugt LEP, et al. Schistosome egg antigens, including the glycoprotein IPSE/alpha-1, trigger the development of regulatory B cells. *PLoS Path.* 2017; 13(7):e1006539. <https://doi.org/10.1371/journal.ppat.1006539> PMID: 28753651
72. Wuhler M, Balog CIA, Catalina MI, Jones FM, Schramm G, Haas H, et al. IPSE/alpha-1, a major secretory glycoprotein antigen from schistosome eggs, expresses the Lewis X motif on core-difucosylated N-glycans. *FEBS J.* 2006; 273(10):2276–92. <https://doi.org/10.1111/j.1742-4658.2006.05242.x> PMID: 16650003
73. Schramm G, Mohrs K, Wodrich M, Doenhoff MJ, Pearce EJ, Haas H, et al. Cutting Edge: IPSE/alpha-1, a Glycoprotein from *Schistosoma mansoni* eggs induces ige-dependent, antigen-independent IL-4 production by murine basophils in vivo. *J Immunol.* 2007; 178(10):6023. PMID: 17475824
74. Elhag SM, Abdelkareem EA, Yousif AS, Frah EA, Mohamed AB. Detection of schistosomiasis antibodies in urine patients as a promising diagnostic maker. *Asian Pac J Trop Dis.* 2011; 4(10):773–7. [https://doi.org/10.1016/S1995-7645\(11\)60192-2](https://doi.org/10.1016/S1995-7645(11)60192-2).
75. Nausch N, Dawson EM, Midzi N, Mduluzi T, Mutapi F, Doenhoff MJ. Field evaluation of a new antibody-based diagnostic for *Schistosoma haematobium* and *S. mansoni* at the point-of-care in northeast Zimbabwe. *BMC Infect Dis.* 2014; 14(1):165. <https://doi.org/10.1186/1471-2334-14-165> PMID: 24666689
76. Milner T, Reilly L, Nausch N, Midzi N, Mduluzi T, Maizels R, et al. Circulating cytokine levels and antibody responses to human *Schistosoma haematobium*: IL-5 and IL-10 levels depend upon age and infection status. *Parasit Immunol.* 2010; 32(11-12):710–21. <https://doi.org/10.1111/j.1365-3024.2010.01235.x> PMID: 21039611
77. Song HB, Kim J, Jin Y, Lee JS, Jeoung HG, Lee YH, et al. Comparison of ELISA and urine microscopy for diagnosis of *Schistosoma haematobium* infection. *J Korean Med Sci.* 2018; 33(33).
78. Itoh M, Ohta N, Kanazawa T, Nakajima Y, Minai M, Daren Z, et al. Southeast Asian J Trop Med Public Health 2003, 34(3) 469–472, A tool for surveillance of schistosomiasis japonica. 1996;3:469–72. PMID: 15115114
79. Midzi N, Mduluzi T, Chimbari MJ, Tshuma C, Charimari L, Mhlunga G, et al. Distribution of schistosomiasis and soil transmitted helminthiasis in Zimbabwe: towards a national plan of action for control and elimination. *PLoS Negl Trop Dis.* 2014; 8(8). <https://doi.org/10.1371/journal.pntd.0003014> PMID: 25121489
80. Katz N, Chaves A, Pellegrino J. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. *Rev Inst Med Trop Sao Paulo.* 1972; 14(6):397–400. PMID: 4675644.
81. World Health Organization. Prevention and control of schistosomiasis and soil-transmitted helminthiasis: World Health Organization/Unicef joint statement. Geneva: World Health Organization <http://www.who.int/iris/handle/10665/68607>. 2004.
82. Corstjens PL, De Dood CJ, Kornelis D, Fat EM, Wilson RA, Kariuki TM, et al. Tools for diagnosis, monitoring and screening of *Schistosoma* infections utilizing lateral-flow based assays and upconverting phosphor labels. *Parasitology.* 2014; 141(14):1841–55. Epub 2014/06/17. <https://doi.org/10.1017/S0031182014000626> PMID: 24932595; PubMed Central PMCID: PMC4265670.
83. Tucker MS, Karunaratne LB, Lewis FA, Freitas TC, Liang Y-s. Schistosomiasis. *Curr Protoc Immunol.* 2013; 103(1):19.1.1–1.58. <https://doi.org/10.1002/0471142735.im1901s103> PMID: 24510597
84. Dalton JP, Day SR, Drew AC, Brindley PJ. A method for the isolation of schistosome eggs and miracidia free of contaminating host tissues. *Parasitology.* 1997; 115 (Pt 1):29–32. Epub 1997/07/01. PMID: 9226954.
85. Basch PF. Cultivation of *Schistosoma mansoni* in vitro. I. Establishment of cultures from cercariae and development until pairing. *J Parasitol.* 1981; 67(2):179–85. Epub 1981/04/01. PMID: 7241277.
86. Roberts SM, MacGregor AN, Vojvodic M, Wells E, Crabtree JE, Wilson RA. Tegument surface membranes of adult *Schistosoma mansoni*: development of a method for their isolation. *Mol Biochem Parasitol.* 1983; 9(2):105–27. Epub 1983/10/01. PMID: 6669162.
87. Vaudel M, Barsnes H, Berven FS, Sickmann A, Martens L. SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics.* 2011; 11(5):996–9. Epub 2011/02/22. <https://doi.org/10.1002/pmic.201000595> PMID: 21337703.
88. Shteynberg D, Nesvizhskii AI, Moritz RL, Deutsch EW. Combining results of multiple search engines in proteomics. *Mol Cell Proteomics.* 2013; 12(9):2383–93. Epub 2013/05/31. <https://doi.org/10.1074/mcp.R113.027797> PMID: 23720762; PubMed Central PMCID: PMC3769318.
89. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics.* 2004; 20(9):1466–7. Epub 2004/02/21. <https://doi.org/10.1093/bioinformatics/bth092> PMID: 14976030.

90. Kim S, Pevzner PA. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun.* 2014; 5:5277. Epub 2014/11/02. <https://doi.org/10.1038/ncomms6277> PMID: 25358478; PubMed Central PMCID: PMC45036525.
91. Eng JK, Jahan TA, Hoopmann MR. Comet: an open-source MS/MS sequence database search tool. *Proteomics.* 2013; 13(1):22–4. Epub 2012/11/14. <https://doi.org/10.1002/pmic.201200439> PMID: 23148064.
92. Tabb DL, Fernando CG, Chambers MC. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res.* 2007; 6(2):654–61. Epub 2007/02/03. <https://doi.org/10.1021/pr0604054> PMID: 17269722; PubMed Central PMCID: PMC15255619.
93. Diament BJ, Noble WS. Faster SEQUEST searching for peptide identification from tandem mass spectra. *J Proteome Res.* 2011; 10(9):3871–9. Epub 2011/07/19. <https://doi.org/10.1021/pr101196n> PMID: 21761931; PubMed Central PMCID: PMC3166376.
94. Vaudel M, Burkhardt JM, Zahedi RP, Oveland E, Berven FS, Sickmann A, et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotechnol.* 2015; 33(1):22–4. Epub 2015/01/13. <https://doi.org/10.1038/nbt.3109> PMID: 25574629.
95. Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM, et al. The PRoteomics IDentification (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 2013; 41(Database issue):D1063–9. Epub 2012/12/04. <https://doi.org/10.1093/nar/gks1262> PMID: 23203882; PubMed Central PMCID: PMC3531176.
96. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 2011; 39(Web Server issue):W29–W37. <https://doi.org/10.1093/nar/gkr367> PubMed PMID: PMC3125773. PMID: 21593126
97. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics.* 2017; 33(18):2938–40. Epub 2017/06/25. <https://doi.org/10.1093/bioinformatics/btx364> PMID: 28645171; PubMed Central PMCID: PMC5870712.
98. Greiner M, Pfeiffer D, Smith RD. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med.* 2000; 45(1–2):23–41. Epub 2000/05/10. PMID: 10802332.